

HIGH-QUALITY, REDUCED DATA RATE STREAMING VIDEO PRODUCTION AND MONITORING SYSTEM

REFERENCE TO RELATED APPLICATION

This application claims priority from U.S. Provisional Patent Application Serial No. 60/411,474, filed September 17, 2002, the entire content of which is incorporated herein by reference.

5

FIELD OF THE INVENTION

This invention relates generally to digital video and, more particularly, to a multi-format digital video production system capable of maintaining full-bandwidth resolution while providing professional quality editing and manipulation of images for various applications, including digital HDTV and specialized video monitoring.

10

BACKGROUND OF THE INVENTION

Traditional systems for video production either rely on uncompressed video signals (for example, SMPTE 4:4:4 or 4:2:2), standard compressed MPEG-2 4:2:2P@ML signals, or other signals that have undergone only minimal compression, such as the (approximately) 5:1 compression utilized for DVCPRO and DVCAM equipment by
15 Panasonic and Sony. However, the bandwidth required for these high-quality signals still is too great for many broadcast and industrial applications, particularly those that require the level of detail available in HDTV images.

Due to the high-bandwidth demands of high-quality signals, typical distribution systems utilize only the highest quality levels for the head-end equipment and the first
20 part of the signal distribution chain. Furthermore, because of network traffic due to multiple users (as for example, in a cable television distribution system), the last leg of the signal path utilizes a more highly compressed signal, to maximize the usage of the available bandwidth. In most cases, this requires that the original signal be decompressed,

and then re-compressed at a much higher compression ratio, so that less bandwidth is required for the final portion of the path.

Figure 1 is a diagram which illustrates the way in which conventional broadband transmission media area used, showing how high compression and interlaced video are poor choices for higher quality applications. Interlaced video, whether NTSC or PAL, is inferior for Internet and broadband applications, since the delivered video quality is less than that possible with progressive display, regardless of compression. Even using a progressive format, however, film (35-mm) and high-end 24 fps progressive camera inputs are deficient in terms of quality, due to the need for high levels of compression later in the signal transmission path.

Accordingly, the need remains for an approach to video production and monitoring which allows the levels of quality that users have come to expect at their receiving terminals, while utilizing existing broadband media and other conventional technologies to optimize the signal storage, processing, and transmission path performance.

SUMMARY OF THE INVENTION

This invention resides in a multi-format digital video production system capable of maintaining the full-bandwidth resolution of the subject material, while providing professional quality editing and manipulation of images intended for digital television and for other applications, including digital HDTV programs and specialized video monitoring applications.

Broadly, this invention allows emerging broadband video transmission media, including Internet broadcast schemes, to overcome existing technology limitations. In the preferred embodiment, for example, the approach facilitates high-quality/large-screen video production and monitoring through the use of conventional broadband channels, including those which currently only exhibit bandwidths on the order of 4 Mbps. In more specific examples, in formats utilizing a 24 fps progressive scan multi-format system, direct streaming is made possible from HDTV (16:9) high-quality data, thereby

expanding market applications which require these higher levels of resolution, bits per pixel, and so forth.

This system, now known as the "Direct Stream Cinema System," is based on optimizing the entire signal path, utilizing 4:2:2 color processing and bit rates typically in the range of 2-6 Mbps. It begins with digitizing and compressing the output of the optical pickup and graphics processor (including appropriate processing, such as noise reduction and resolution enhancement) and carries through the processing circuitry to the receiving terminal device at the user end of the transmission chain. Signal quality is preserved throughout the process, by eliminating the need to decompress a lower-compression signal from a camera, video recorder, or other source device for editing or other purposes, and then re-compressing the signal at a much higher rate for transmission purposes.

A high-quality, reduced-data-rate digital video system according to a preferred embodiment includes a source of a streaming video program having a progressive-scanned image with a frame rate of less than substantially 24 fps; a video server in communication with the source for storing the program; and one or more computers in network communication with the video server for locally displaying the program or portions thereof.

In a "direct stream" implementation the locally displayed program or portions thereof are in the same format as the streaming video program received from the source. The system and method may further include a personal-computer-based control of the camera/input device, monitor for the streaming video program received from the source, or other PC-based capabilities. The streaming video program may be received through a network connection, and the video server includes one or more of the following for storing the program: a micro-disk, portable HDD, memory-stick, optical storage, or magneto-optical storage.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1 is a diagram which illustrates the way in which conventional broadband transmission media are used, showing how high compression and interlaced video are poor choices for higher quality applications;

FIGURE 2 is a diagram which shows the concepts behind the various versions of the "Direct Stream Cinema" systems;

FIGURE 3 is a diagram which illustrates a particular configuration constructed in accordance with the invention, namely, a video surveillance system;

FIGURE 4 is a drawing which shows a different particular layout according to the invention, in this a streaming production system;

FIGURE 5 is a diagram which shows the way in which the instant invention implements HDTV quality video at a very low overall system cost; and

FIGURE 6 is a diagram which shows the quality levels provided by the consumer-type implementation of the "Direct Stream Cinema" systems.

DETAILED DESCRIPTION OF THE INVENTION

This invention overcomes the limitations of the existing art by providing a video production/monitoring capability capable of transmitting HDTV (16:9) quality video utilizing existing broadband bandwidths of [such as 4 Mbps (1024 x 576 pixels) or greater], thereby overcoming the traditional problem of conserving bandwidth while preserving quality.

Referring back to Figure 1, the "Direct Stream Cinema System" preferably utilizes a 24 fps progressive camera format which, through the use of proprietary multi-format production techniques, facilitates Internet and broadband applications, including streaming services, Internet TV, video monitoring/security, and 35 mm/HDTV/DVD output capabilities. The approach does not require an HDTV quality video camera or recording, however, but nevertheless facilitates HDTV quality, direct video monitoring, off-line editing, and other capabilities at a great reduction in total system cost.

With respect to streaming applications, the video data may be transmitted directly to a central server through a network environment, resulting in both a comparatively small capacity storage requirement and also in other advantages over existing approaches. In one disclosed example, HDTV quality video with an aspect ratio of 16:9 is achieved, 5 having a horizontal resolution of 1024 x 576, with the potential for up-conversion to 1920 x 1080. This resolution, equivalent to a 42-inch plasma display, is accomplished with a data rate of 4 Mbps, more or less, enabling recording to occur at 2 GBytes/hr, whereas current HDTV requires more than 100 GBytes/hr. Various video formats are possible through the use of proprietary multi-format progressive systems and frame rates, which 10 may vary up to 24 fps (or greater) in the preferred embodiment.

Newer media players, such as Microsoft's new "Corona" technology, which is scheduled to be released with the latest version of the Windows Media Player (Series 9), are aimed at signal distribution systems utilizing a data rate of 6 Mbps, using MPEG 4 and other comparable compression techniques. However, such technology also provides 15 for bit rates in the range of 2-4 Mbps, being directed towards applications such as archiving, streaming video, and off-line viewing. At these data rates, it is possible to store 100 hours of video in only 180 GB of storage $[(100 \text{ hr}) \times (3600 \text{ sec/hr}) \times (4 \text{ Mbps}) / (8 \text{ b/B})]$.

Figure 2 discloses three of the many potential implementations of the "Direct 20 Stream Cinema" system: Professional cameras and Camcorders, Consumer Camcorders, and Digital-Still-Camera recorders. Video stored in this manner is suitable for on-line editing applications, using PC plug-in hardware cards from companies like Matrox (Perphelia) and ATI (Radion 9200/9800), Nvidia (GeForce FX). However, these conventional off-the-shelf-types of cards require modification, so that they would be 25 equipped with true DV or SDI digital video outputs, thereby providing compatibility with both HDTV and standard NTSC formats, including analog, Y-C component formats, and composite video outputs. In addition, software packages such as Adobe Premier 6.5, and Ulead MediaStudio 7, when utilized with a high-end PC (3 GHz or higher processing speed), are capable of providing sophisticated editing capabilities.

The resulting signal can be stored, in an AVI format, for example, on a hard disk drive. Currently, these PC cards only are being used for SDTV, but in the future, they will be capable of HDTV recording, and for specialized industrial applications; for HDTV applications, a new decoder board would be used.

5 The preferred storage and distribution format according to the invention is 1024 x 576 @ 24 fps. Compression ratios of 100:1 are practical for SDTV, and 400:1 for HDTV. In addition, the system is scalable, for example, to the following:

 200 Kbps @ 1 fps
 500 Kbps @ 3 fps
10 1 Mbps @ 6 fps
 2 Mbps @ 12 fps
 4 Mbps @ 24 fps

 Comparisons of the output quality of a variety of PC-video display cards utilizing both interlaced and progressive signals and also frame-rate/standards-conversion indicate
15 a need to optimize the signal processing. For conversions from interlaced PAL signals to NTSC, these cards produce outputs with noticeable frame skipping and jumping. However, from a progressive PAL signal (i.e., greater than 50 fps progressive), the severity of artifacts is greatly reduced. Newer PC graphics cards produce significantly better results, which suggests that they may have adopted the frame-rate conversion
20 techniques disclosed in U.S. Patent No. 5,999,220, entitled "Multi-Format Audio/Video Production System with Frame Rate Conversion" and U.S. Patent No. 6,370,198B1, entitled "Wide-band Multi-Format Audio/Video Production System with Frame Rate Conversion," the entire content of both being incorporated herein by reference.

 In preferred embodiments, signals at the head-end of a signal distribution system
25 are converted to progressively scanned signals. A frame rate of 24 fps preferably is employed, in order to optimize the utilization of the available bandwidth. In the next step, the signals are compressed to create a data stream at 2-4 Mbps (for 1024 x 576 @ 24 fps) or 4-6 Mbps (for 1280 x 720 @ 24 fps. These signals may be stored for subsequent transmission to receiving terminal equipment (such as PCs, cable boxes, personal video

recorders, display monitors, or other terminal equipment), or immediately transmitted over a signal distribution system, which may be wired, wireless, satellite, or other medium, including physical media such as CD-ROMs, DVDs, etc.). This receiving terminal equipment may be located at multiple remote sites, may be located at multiple sites within a single facility, or may be configured as a combination of local and remote sites.

In an alternative embodiment, signals may be received from multiple sources, including one or more remote sources, and are collected at a central location for viewing, storage, or both. The signals preferably are transmitted to the central site as compressed, progressively-scanned streaming video signals, employing data rates in the range of 2-4 Mbps. As in other embodiments, 24 fps is preferably used, although the frame rate may be greater or less, may be variable or fixed, and may be modified under control of a local operator, or may be modified automatically in response to a predetermined set of criteria, utilizing sensors at the physical location of the camera or signal source, or via remote control from a central site, either under control of an operator, or automatically in response to a predetermined set of criteria. The source signal frame rate and image size may be different for each source signal, and the frame rate and image size of a source signal in the format stored need not be identical to the frame rate and image size in the format displayed.

Currently, 1/2-inch 3-CCD cameras are available for less than \$10,000, and 1/3-inch 3-CCD cameras are available for approximately \$5,000. As such, it is already practical and economical to implement this type of system for a range of commercial/industrial applications, for example:

- Airport security
- Monitoring of remote natural areas, such as forests
- Auto crash testing
- Public building (Court, Government office, School) security
- Hospital security
- Educational/instructional

Figure 23 is a diagram which illustrates a particular configuration constructed in accordance with the invention, namely, a video surveillance system. In this case, the signals from multiple cameras are transmitted as streaming sources at relatively low data rates, on the order of 200K to 4 Mbps, with 1 to 24 fps variable frame rates. As discussed
5 above, this conserves video server storage requirements, facilitating one hour of storage utilizing only 2 GB of capacity. This information may then be network-accessed by one or more monitoring systems, preferably using multi-screen displays, and optionally including alarms or other features using graphic analysis or other methodologies.

Figure 4 is a drawing which shows a different particular layout according to the
10 invention, in this case a streaming production system which may be implemented with Professional-quality equipment. Again, a camera producing HDTV quality video transmits at a relatively low data rate as a streaming source to a program editing facility, enabling various operator controls including, but not limited to, frame-by-frame control, variable playback, forward/reverse (bi-directional) playback, and so forth. A decision list
15 is generated on a scene-by-scene basis, with AVI file conversion being used for compatibility with PC non-linear editing. Alternative formats would include, for example, MPEG-4, Windows Media 9, or Divx (which even can be edited, utilizing one of the available software packages for editing. The source material and EDL (Edit Decision List) codes are stored in a streaming server, with the resulting modest
20 requirements facilitating an hour of storage within a Gigabyte of memory (for SDTV at 2 Mbps) or within two Gigabytes of memory (for HDTV at 4 Mbps. The streaming video is output to one or more likely multiple viewing stations, utilizing an even lower data rate of, perhaps, less than two Mbps. Conventional SDTV signals utilizing a compressed DV-type output typically would be provided at 25-50 Mbps. HDTV-type signals utilizing a
25 compressed SDI-type output would be provided at 100-300 Mbps; however, the signal manipulations within the system and before the output stages would utilize the more efficient and compact 4 Mbps files and signal streams.

Figure 5 is a diagram which shows the way in which the instant invention implements HDTV quality video at a very low overall system cost. At the high end, an

HDTV camera with a format of 16:9 at 1920 x 1080 pixels uses some 2 million pixels per image as the source, which is reduced to less than 1 Megapixels or thereabouts due to interlace losses, bandwidth limiting, compression losses and so forth. Even so, equipment exhibiting this level of performance currently involves hardware costs of approximately \$200,000.

While broadcast quality video (standard definition at 4:3) costs much less, the image quality is reduced dramatically, to a frame size of 720 x 480 pixels. According to the invention, however, utilizing a 24 fps scan and proprietary multi-format system, an image at 1024 x 576 can be generated having an aspect ratio of 16:9, exhibiting a quality comparable to conventional HDTV broadcast, but at a cost of under \$10,000. A typical surveillance image, at 320 x 240 is shown for comparison purposes.

For any of these implementations (Professional, Camcorder, Surveillance, or Consumer), a key part of the system resides in the optimization of the entire processing scheme, with an eye towards the end-user quality level. For example, in the case of modern plasma-display units, the capability of the individual unit largely is determined by the physical dimensions of the screen: 32" displays are supplied as capable of 848 x 477 pixels; 42" displays are supplied as capable of 1024 x 576 pixels; 50" displays are supplied as capable of 1280 x 720 pixels. Because multiple tests have demonstrated that "film quality" as measured at the theatrical projection screen only provides approximately 700 lines of resolution (see, for example, A. Kaiser, H. W. Mahler, and R. H. McMann, *SMPTE Journal*, June, 1985), 1024 x 576, or at most 1280 x 720, provides the optimum display quality; 1920 x 1080 or other higher-pixel-count systems are not required.

Another key feature of the system is the utilization of compression technology. Most origination-quality systems rely on intra-frame compression (such as Motion-JPEG), which is limited to 3:1 or 4:1 for this type of application. Further downstream in the processing and transmission chain, much higher inter-frame-based compression ratios are needed (such as MPEG-2), in order to make signal distribution practical and economical. The instant invention contemplates high compression ratios throughout the

process, achieving in excess of 100:1 compression. In this way, the use of "intermediate" formats, such as DVC-PRO or DV-CAM no longer are required. Furthermore, the reduced data rates required for the system eliminates the need for extremely large capacity hard-disk recording capability, enabling editing on most of today's conventional
5 PCs.

However, in order to achieve these kinds of compression ratios without sacrificing quality, the preferred embodiment employs 24 fps signals (which, evidently, saves 20% of the data rate required for a 30 fps signal), and also progressive-scanning (which is over 50% more efficient than compression of interlaced signals). Many compression schemes
10 are available, including, for example, industry standards such as MPEG-4, and proprietary systems such as Microsoft Windows Media 9, Divx, and Wavelet-type compression. The resulting data rates easily are conveyed over conventional distribution paths, such as satellite, cable, and broadcast systems, requiring only 1-2 Mbps for SDTV-type signals, and 6 Mbps for HDTV-type signals.

15 As shown in Figure 6, in Consumer-type applications, it is common to employ digital still camera systems, utilizing high-speed shutters to provide video program sourcing. The results are limited to relatively low-quality recordings for relatively limited recording times. In addition, many artifacts are imparted to the recordings, such as motion artifacts and picture hesitation or jumps. By employing the techniques
20 disclosed herein, DV-quality recordings for more than one hour are practical, and S-VHS-quality recordings for more than two hours can be achieved. In addition, video editing is simplified, as no step of capturing to the PC is required -- editing can proceed directly from camera memory cards or other storage devices (including hard-disk, optical disc, DVD, *etc.*), and the quality is preserved throughout the process. In addition, the
25 resulting recordings are compatible with various streaming conventions, such as those supported by Microsoft and Real Networks video. This same system of video processing without a step of capturing the signal to the PC applies equally as well to Professional and Camcorder applications.

The reader will appreciate that the practical application of the instant invention has significant implications in many fields. For example, Digital Asset Management systems typically employ highly-compressed "proxies" to convey the content of much less-compressed primary program materials, thereby enabling Edit Decision Lists to be developed from the "proxies" and then used to edit the final program using the primary program material. With the much more efficient signal processing methods provided herein, it is not necessary to create the separate proxies, as the primary signals themselves are provided at much lower data rates than traditionally have been available for these materials, making them suitable for use in a single-step on-line editing application.

10 The "Direct Stream Cinema System" is based on optimizing the entire signal path, utilizing 4:2:2 color processing and bit rates typically in the range of 1-2 Mbps for SDTV-quality video and 4-6 Mbps for HDTV-quality video. It begins with digitizing and compressing the output of the optical pickup and graphics processor (including appropriate processing, such as noise reduction and resolution enhancement), so that
15 from the onset the data rate is set and then maintained through the internal processing circuitry, recording steps, and through the distribution steps to the receiving terminal device at the user end of the transmission chain. Signal quality is preserved throughout the process, by eliminating the need to decompress a lower-compression signal from a camera, video recorder, or other source device for editing or other purposes, and then re-
20 compressing the signal at a much higher rate for transmission purposes. Thus, there is no distinct "intermediate" format of any kind, as the original video format obtained from the optical pickup or other source device is maintained through the entire path to the receiving terminal device.

 Note that, to a certain extent, the resolution sizes and pixels, as well as the prices,
25 and other data are associated with current technology, and are anticipated to vary in time as technology improves and matures. Nevertheless, the inventive approach of applicant will at all times result in a substantial decrease in system cost while preserving the highest possible quality, even at limited bandwidths. Additionally, in all embodiments of

FNI-02902/03
31709sh

the invention, techniques such as pixel interpolation may advantageously be used to further enhance image resolution/quality.

I claim: